X-ray characterization of PaPheOH, a bacterial phenylalanine hydroxylase

AKADEMISK AVHANDLING

Som med vederbörligt tillstånd av rektorsämbetet vid Umeå Universitet för avläggande av filosofie doktorsexamen framläggs till offentligt försvar i sal D, byggnad 1D, fredagen den 3 oktober kl. 10.00.

av

Fredrik Ekström
Umeå Centrum för Molekylär Patogenes
Umeå Universitet

Fakultetsopponent: Prof. Edward Hough, Department of Chemistry, University of Tromsø, Norge.
Many human diseases are associated with the malfunction of enzymes in the aromatic amino acid hydroxylase family, e.g. phenylketonuria (PKU), hyperphenylalaninemia (HPA), schizophrenia and Parkinson's disease. The family of aromatic aminoacid hydroxylases comprises the structurally and functionally related enzymes phenylalanine hydroxylase (PheOH), tyrosine hydroxylase (TyrOH) and tryptophane hydroxylase (TrpOH). These enzymes require the cofactor (6R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH$_4$) and atomic oxygen. In eukaryotes, the aromatic amino acid hydroxylases share the same organization with a N-terminal regulatory domain, a central catalytic domain and a C-terminal tetramerization domain. Aromatic amino acid hydroxylases that correspond to the core catalytic domain of the eukaryotic enzymes are found in bacteria. The main focus of this thesis is the structural characterization of a phenylalanine hydroxylase from the opportunistic pathogen Pseudomonas aeruginosa (PaPheOH).

In order to investigate the active site environment and to probe the oxidation state of the active site iron X-ray absorption spectroscopy (XAS) experiments were initiated. The experimental data support a model where the active site iron is coordinated by four oxygen atoms and two nitrogen atoms. We suggest that two water molecules, His121, His126 and Glu166 coordinates the active site iron. In this model, Glu166 provides two of the oxygen atoms in a bidentate binding geometry. EXAFS and XANES studies indicate that the iron is in a ferrous (Fe(II)) state and that no structural rearrangements are induced in the first coordination shell in samples of PaPheOH with BH$_4$ and/or L-phenylalanine.

The 1.6 Å X-ray structure of PaPheOH shows a catalytic core that is composed of helices and strands in a bowl-like arrangement. The iron is octahedrally coordinated, by two water molecules and the evolutionary conserved His121, His126 and Glu166 that coordinates the iron with bidentate geometry. The pterin binding loop of PaPheOH (residue 81-86) adopts a conformation that is displaced by 5-6 Å from the expected pterin binding site. Consistent with the unfavourable position of the pterin binding loop is the observation that PaPheOH has a low specific activity compared to the enzymes from human and Chromobacterium violaceum.

The second part of this thesis focus on the crystallization and structure determination of the actin binding domain of α-actinin (ABD). α-Actinin is located in the Z-disc of skeletal muscle were it crosslinks actin filaments to the giant filamentous protein titin. The ABD domain of α-actinin crystallizes in space group P2$_1$ with four molecules in the asymmetric unit. The structure of the ABD domain has been solved to a d-spacing of 2.0 Å. The two CH-domains of ABD are each composed of 5 α-helices. The α-helices pack into a closed compact conformation with extensive intramolecular contacts between the two domains.

Keywords: PaPheOH, PheOH, PAH, phhA, phenylalanine hydroxylase, protein crystallography, α-actinin, microcrystal, ABD, CH-domain.
X-ray characterization of PaPheOH, a bacterial phenylalanine hydroxylase

by

Fredrik Ekström
Denna avhandling tillägnas mina föräldrar
Margot och Bernt Ekström
ABSTRACT

MAIN REFERENCES

1. AROMATIC AMINO ACID HYDROXYLASES

1.1 THE AROMATIC AMINO ACID HYDROXYLASES
1.2 OVERALL STRUCTURE OF THE AROMATIC AMINO ACID HYDROXYLASES
1.3 THE ENZYMATIC REACTION OF THE AROMATIC AMINO ACID HYDROXYLASES
1.4 COFACTOR PRODUCTION AND RECYCLING
1.5 CRYSTALLOGRAPHIC STUDIES OF AROMATIC AMINO ACID HYDROXYLASES
1.6 THE ACTIVE SITE
1.7 ENZYMATIC MECHANISM OF HUMAN PHENYLALANINE HYDROXYLASE

2. THE GRAM-NEGATIVE, OPPORTUNISTIC PATHOGEN PSEUDOMONAS AERUGINOSA

2.1 THE AROMATIC AMINO ACID HYDROXYLATING SYSTEM OF PSEUDOMONAS AERUGINOSA
2.2 PAPCD/PHHB, A BACTERIAL ORTHOLOGUE OF PCD/DCOH

3. AIMS OF THE PRESENT STUDY

4. X-RAY CRYSTALLOGRAPHY

4.1 WHY WE USE X-RAYS AND CRYSTALS
4.2 CRYSTALLIZATION OF PROTEINS
4.3 DATA COLLECTION AND ANALYSIS
4.4 THE INITIAL MODEL
4.5 MODEL BUILDING AND STRUCTURE REFINEMENT

5. X-RAY ABSORPTION SPECTROSCOPY

5.1 X-RAY ABSORPTION SPECTROSCOPY
5.2 INTERPRETATION OF XAS SPECTRA

6. PAPER I

7. PAPER II

8. PAPER III

9. PAPER IV

10. THE 2.0 Å X-RAY STRUCTURE OF THE ACTIN BINDING DOMAIN OF α-ACTININ

11. CONCLUSIONS

13. ACKNOWLEDGEMENTS

14. REFERENCES
Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABD</td>
<td>Actin binding domain</td>
</tr>
<tr>
<td>BH₄</td>
<td>(6R)-L-erythro-5,6,7,8-tetrahydrobiopterin</td>
</tr>
<tr>
<td>7,8-BH₂</td>
<td>L-erythro-7,8-dihydrobiopterin</td>
</tr>
<tr>
<td>CH-domain</td>
<td>Calponin Homology domain</td>
</tr>
<tr>
<td>CvPheOH</td>
<td>Chromobacterium violaceum phenylalanine hydroxylase</td>
</tr>
<tr>
<td>ESRF</td>
<td>European synchrotron radiation facility</td>
</tr>
<tr>
<td>EXAFS</td>
<td>Extended X-ray absorption fine structure</td>
</tr>
<tr>
<td>HPA</td>
<td>Hyperphenylalaninemia</td>
</tr>
<tr>
<td>hPheOH</td>
<td>Human phenylalanine hydroxylase</td>
</tr>
<tr>
<td>PAH</td>
<td>Phenylalanine hydroxylase</td>
</tr>
<tr>
<td>PaPheOH</td>
<td>Pseudomonas aeruginosa phenylalanine hydroxylase</td>
</tr>
<tr>
<td>PheOH</td>
<td>Phenylalanine hydroxylase</td>
</tr>
<tr>
<td>PKU</td>
<td>Phenylketonuria</td>
</tr>
<tr>
<td>rPheOH</td>
<td>Rattus norvegicus phenylalanine hydroxylase</td>
</tr>
<tr>
<td>rTyrOH</td>
<td>Rattus norvegicus tyrosine hydroxylase</td>
</tr>
<tr>
<td>Se-Met</td>
<td>Seleno-L-methionine</td>
</tr>
<tr>
<td>MAD</td>
<td>Multiple anomalous dispersion</td>
</tr>
<tr>
<td>MIR</td>
<td>Multiple isomorphous replacement</td>
</tr>
<tr>
<td>MR</td>
<td>Molecular replacement</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>SAD</td>
<td>Single wavelength anomalous dispersion</td>
</tr>
<tr>
<td>TyrOH</td>
<td>Tyrosine hydroxylase</td>
</tr>
<tr>
<td>TrpOH</td>
<td>Tryptophane hydroxylase</td>
</tr>
<tr>
<td>XANES</td>
<td>X-ray absorption near edge structure</td>
</tr>
<tr>
<td>XAS</td>
<td>X-ray absorption spectroscopy</td>
</tr>
</tbody>
</table>
Abstract

Many human diseases are associated with the malfunction of enzymes in the aromatic amino acid hydroxylase family, e.g. phenylketonuria (PKU), hyperphenylalaninemia (HPA), schizophrenia and Parkinson's disease. The family of aromatic amino acid hydroxylases comprises the structurally and functionally related enzymes phenylalanine hydroxylase (PheOH), tyrosine hydroxylase (TyrOH) and tryptophane hydroxylase (TrpOH). These enzymes require the cofactor (6R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH$_4$) and atomic oxygen. In eukaryotes, the aromatic amino acid hydroxylases share the same organization with a N-terminal regulatory domain, a central catalytic domain and a C-terminal tetramerization domain. Aromatic amino acid hydroxylases that correspond to the core catalytic domain of the eukaryotic enzymes are found in bacteria. The main focus of this thesis is the structural characterization of a phenylalanine hydroxylase from the opportunistic pathogen *Pseudomonas aeruginosa* (PaPheOH).

In order to investigate the active site environment and to probe the oxidation state of the active site iron X-ray absorption spectroscopy (XAS) experiments were initiated. The experimental data support a model where the active site iron is coordinated by four oxygen atoms and two nitrogen atoms. We suggest that two water molecules, His121, His126 and Glu166 coordinates the active site iron. In this model, Glu166 provides two of the oxygen atoms in a bidentate binding geometry. EXAFS and XANES studies indicate that the iron is in a ferrous (Fe(II)) state and that no structural rearrangements are induced in the first coordination shell in samples of PaPheOH with BH$_4$ and/or l-Phe.

The 1.6 Å X-ray structure of PaPheOH shows a catalytic core that is composed of helices and strands in a bowl-like arrangement. The iron is octahedrally coordinated, by two water molecules and the evolutionary conserved His121, His126 and Glu166 that coordinates the iron with bidentate geometry. The pterin binding loop of PaPheOH (residue 81-86) adopts a conformation that is displaced by 5-6 Å from the expected pterin binding site. Consistent with the unfavourable position of the pterin binding loop is the observation that PaPheOH has a low specific activity compared to the enzymes from human and *Chromobacterium violaceum*.

The second part of this thesis focus on the crystallization and structure determination of the actin binding domain of α-actinin (ABD). α-Actinin is located in the Z-disc of skeletal muscle were it crosslinks actin filaments to the giant filamentous protein titin. The ABD domain of α-actinin crystallizes in space group P2$_1$ with four molecules in the asymmetric unit. The structure of the ABD domain has been solved to a d-spacing of 2.0 Å. The two CH-domains of ABD are each composed of 5 α-helices. The α-helices pack into a closed compact conformation with extensive intramolecular contacts between the two domains.

Keywords: PaPheOH, PheOH, PAH, phhA, phenylalanine hydroxylase, protein crystallography, α-actinin, microcrystal, ABD, CH-domain.
Main references

This thesis is based on the following publications and manuscripts, referred to into the text by their roman numerals (I-IV).


1. Aromatic amino acid hydroxylases

The aromatic amino acid hydroxylases catalyze the incorporation of one oxygen atom into the aromatic ring of L-phenylalanine (phenylalanine hydroxylase, PheOH), L-tyrosine (tyrosine hydroxylase, TyrOH) and L-tryptophan (tryptophane hydroxylase, TrpOH). In mammals, this system is important for the production of the neurotransmitters/hormones dopamine, norepinephrine, epinephrine and serotonin. In prokaryotes, aromatic amino acid hydroxylases are involved in the biodegradation and recycling of hydrocarbons.

1.1 The aromatic amino acid hydroxylases
Due to their pivotal role in metabolism, malfunction of the aromatic amino acid hydroxylases are associated with a variety of diseases in humans. As early as 1959, Kaufman related a deficiency in the human enzyme phenylalanine hydroxylase (hPheOH) to the genetic disease phenylketonuria (PKU) (1). Tyrosine hydroxylase (TyrOH) has been implicated in juvenile Parkinsonism (2), L-DOPA responsive dystonia (3), bipolar effective disorder (4), schizophrenia (5) and idiopathic Parkinsonism (Parkinson's disease) (6). Tryptophane hydroxylase (TrpOH) catalyses the first and rate-limiting step in serotonin biosynthesis. Serotonin is involved in numerous physiological functions including sleep, pain, appetite and sexual behaviour. It is also the precursor of the hormone melatonine (7).
1.2 Overall structure of the aromatic amino acid hydroxylases

The human enzymes, PheOH, TyrOH and TrpOH (hPheOH, hTyrOH and hTrpOH respectively) are multi domain proteins composed of a N-terminal regulatory domain (hPheOH 1-142; hTyrOH 1-155; hTrpOH 1-177) and a C-terminal catalytic domain and tetramerization region (hPheOH 143-452; hTyrOH 156-498; hTrpOH 178-445) (8). The tetramerization domain usually comprises the last 20-23 (8-10). The sequences of the catalytic domains of the aromatic amino acid hydroxylases are highly conserved (Figure 1), whereas the regulatory domains are only weakly related. This might reflect their different regulatory mechanisms (11).

The best-characterized prokaryotic aromatic hydroxylases are the phenylalanine hydroxylases from Chromobacterium violaceum (CvPheOH) (12-14) and from Pseudomonas aeruginosa (PaPheOH) (Figure 1) (15, 16). These two enzymes correspond to the catalytic domain of human PheOH with which they share 29% (CvPheOH, 170 residues overlap) and 34% (PaPheOH, 212 residues overlap) sequence identity. The mutual sequence identity between the two bacterial proteins is 50% covering a 131 residues interval. PaPheOH and CvPheOH are monomers in solution since they are lacking the C-terminal tetramerization domain of the eukaryotic proteins. The bacterial proteins also lack the N-terminal regulatory domain suggesting a simpler mode of regulation compared with the eukaryotic family members.
hPheOH  
%  

Figure 1. Sequence alignment of aromatic amino acid hydroxylases. Identical residues are marked with black, the pterin binding loop is marked with □ and residues involved in iron coordination are marked with *.
1.3 The enzymatic reaction of the aromatic amino acid hydroxylases

The catabolic and metabolic task allotted to the aromatic amino acid hydroxylases requires both high selectivity and high conversion rates. The enzymes share the same enzymatic mechanism where one atom of dioxygen is incorporated into the aromatic ring of L-phenylalanine (PheOH), L-tyrosine (TyrOH) or L-tryptophan (TrpOH) (Figure 2). The enzymatic reaction requires the cofactor \((6R)-L\text{-}\text{erythro}-5,6,7,8\text{-}\text{tetrahydrobiopterin} (\text{BH}_4)\) and molecular oxygen (17).

(Figure 2.) Enzymatic reactions catalyzed by the eukaryotic aromatic amino acid hydroxylases. a) Phenylalanine hydroxylase converts L-phenylalanine to L-tyrosine. b) Tyrosine hydroxylase catalyzes the conversion of L-tyrosine to dihydroxyphenylalanine (L-DOPA), a precursor of the neurotransmitters dopamine, norepinephrine and epinephrine. c) Tryptophane hydroxylase catalyses the conversion of L-tryptophan to 5-hydroxy-L-tryptophan, a precursor in the biosynthesis of serotonin and melatonin.
1.4 Cofactor production and recycling

The cofactor (6R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH₄) is required for the enzymatic activity of the aromatic amino acid hydroxylases. The cofactor is synthesized from GTP by GTP cyclohydrolase I (GTPCH; EC 3.5.4.16; PDB entry 1GTP), 6-pyruvoyl tetrahydrobiopterin synthase (6-PTPS; EC 4.6.1.10; PDB entry 1B6Z) and sepiapterin reductase (SR; EC 1.1.1.153; PDB entry 1SEP) (Figure 3).

(Figure 3.) Tetrahydrobiopterin biosynthesis is starting from GTP and is catalyzed by the enzyme GTP cyclohydrolase I (GTPCH), 6-pyruvoyltetrahydrobiopterin synthase (6-PTPS) and sepiapterin reductase (SR).

During the enzymatic reaction of the aromatic hydroxylases the BH₄ molecule is oxidized to pterin-4a-carbinolamid, a molecule that is recycled in a two step reaction by the enzymes pterin-4a-carbinolamine
dehydratase (PCD/DCoH; EC 4.2.1.96; PDB entry 1DCH) and dihydropterin reductase (DHPR; EC 1.6.99.7; PDB entry 1DHR) (Figure 4).

Figure 4. During the enzymatic reaction the tetrahydrobiopterin is oxidized by the aromatic amino acid hydroxylases. The cofactor is then recycled in a two-step reaction by the enzymes pterin-4a-carbinolamid dehydratase (PCD/DCoH) and by the NADH dependent dihydropterin reductase (DHPR).

1.5 Crystallographic studies of aromatic amino acid hydroxylases
During the recent years, several X-ray structures have increased our understanding of this class of enzymes. The most extensively studied enzyme is hPheOH that has been solved in a number of different complexes and constructs. The overall fold of the catalytic domain of hPheOH can be described as a basket-like arrangement of helices, strands and loops (Figure 5a).
Figure 5. a) The catalytic domain of human phenylalanine hydroxylase (hPheOH) with the active site iron shown as a black sphere (N- and C-terminal truncated, residues in model 117-424; PDB entry 1PAH). b) View of the active site. The iron is octahedrally coordinated by His285, His290, Glu330 and three water molecules (shown as grey spheres) (close-up of PDB entry 1PAH). c) The structure of rat phenylalanine hydroxylase with an intact regulatory domain (dark grey) (C-terminally truncated, residues in model 1-429; PDB entry 1PHZ). d) The tetrameric, N-terminally truncated form of hPheOH (residues in model 118-452; PDB entry 2PAH). The figure was prepared using MOLSCRIPT (18).
The active site consists of a deep cleft at the centre of the catalytic domain basket. The active site iron is located at the bottom of the cleft where it is coordinated by the ligands His285, His290, Glu330 and three water molecules (Figure 5b) (19). The structure of rat phenylalanine hydroxylase (rPheOH) has been determined with an intact N-terminal regulatory domain (Figure 5c).

The regulatory domain is composed of a α−β sandwich (βαββαβ) with the N-terminal auto regulatory sequence extending across the active site in the catalytic domain (20). The structure of tetrameric hPheOH shows that the C-terminal tetramerization domain is formed by two β-strands and a 40 Å long α-helix. The C-terminal α-helices from each monomer form a tightly packed anti-parallel coiled-coil motif in the centre of the tetramer (Figure 5d) (9).

The structure of hPheOH in complex with the cofactor analogue 7,8-dihydrobiopterin (hPheOH-Fe(III)·7,8-BH$_2$) demonstrated that the cofactor binds close to, but without any direct contact to the active site iron. The structure also demonstrated that the pterin binding loop between residues 245 and 250 moves towards the iron upon cofactor binding. This allows important hydrogen bonding interactions between the loop and the BH$_4$ cofactor. A π-stacking interaction between Phe254 and the ring system of the pterin cofactor and a water mediated hydrogen bound between Glu286 and the N3 atom of the pterin ring were identified as crucial for the enzymatic process (21).

The first structure of an aromatic aminoacid hydroxylase in its reduced form (hPheOH-Fe(II)) and in complex with its natural cofactor (hPheOH-Fe(II)·BH$_4$) was presented by Andersen et al. (22). This was
followed by the structure determination of reduced hPheOH in complex with the substrate analogue 3-(2-thienyl)-L-alanine (hPheOH-Fe(II)-BH$_4$·THA). In the ternary structure of hPheOH, the global structural changes induced by substrate binding were observed (23). It was shown that Glu330 changes from monodentate to bidentate iron coordination upon substrate binding (carboxylate shift). The bidentate iron coordination has been proposed to be important for oxygen activation (23).

A recent X-ray structure of a phenylalanine hydroxylase from the bacterium *Chromobacterium violaceum* (CvPheOH) has provided a structural explanation for the high specific activity of this bacterial enzyme (14). Similar to the hPheOH-Fe(II)-BH$_4$·THA structure, the active site glutamic acid (Glu184) of CvPheOH-Fe(III) and CvPheOH-Fe(III)-7,8-BH$_2$ coordinates the iron in a bidentate fashion. In the same study, the first structure of a phenylalanine hydroxylase in its apo form (with no active site iron) was reported (14).

The binding of catechol inhibitors to the active site of hPheOH has also been investigated by X-ray crystallography. The study demonstrated that the inhibitors occupy the cofactor-binding site thereby interfering with catalysis (24).

In addition to the crystallographic studies of various forms of PheOH, the structure of TyrOH (25, 26) and TrpOH has been determined (7). As expected from sequence alignments, the catalytic domains of PheOH, TyrOH and TrpOH share the same fold.
1.6 The active site

The active site of the aromatic amino acid hydroxylases is composed of a 2-His-1-carboxylate motif binding a metal ion (27). Additional water molecules to obtain the octahedral 6 coordination, which is the most frequently observed coordination for these enzymes, further coordinate the metal. The active site metal is usually iron, but a copper dependent phenylalanine hydroxylase has been reported from *Chromobacterium violaceum* (CvPheOH-Cu) (28, 29). However, recent studies of CvPheOH (13) and the crystal structure (14) of CvPheOH report an iron at the active site. These conflicting reports could be due to differences in the sequence between residue Leu172 and Ala273 (14, 30). In the aromatic amino acid hydroxylases the carboxylate residue of the 2-His-1-carboxylate motif is a glutamic acid binding in monodentate or bidentate geometry. The geometry of the glutamic acid and the number of coordinated water molecules within this arrangement varies in the available X-ray structures, summarized in Table 1.
Table 1. Summary of selected X-ray structures.

<table>
<thead>
<tr>
<th>Structure</th>
<th>PDB identifier</th>
<th>No. waters</th>
<th>Glu binding distances (Å)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>hPheOH-Fe(III)</td>
<td>1PAH</td>
<td>3</td>
<td>2.1/3.5 (Glu330)</td>
<td>(19)</td>
</tr>
<tr>
<td>hPheOH-Fe(II)</td>
<td>1J8T</td>
<td>2</td>
<td>2.4/4.2 (Glu330)</td>
<td>(22)</td>
</tr>
<tr>
<td>hPheOH-Fe(II)·BH₄</td>
<td>1J8U</td>
<td>3</td>
<td>2.1/3.2 (Glu330)</td>
<td>(22)</td>
</tr>
<tr>
<td>hPheOH-Fe(II)·BH₄·THA</td>
<td>1KW0</td>
<td>1</td>
<td>2.4/2.6 (Glu330)</td>
<td>(23)</td>
</tr>
<tr>
<td>hPheOH-Fe(II)·7,8-BH₂</td>
<td>1DMW</td>
<td>3</td>
<td>2.0/3.3 (Glu330)</td>
<td>(21)</td>
</tr>
<tr>
<td>rPheOH-Fe(III)-Ser16-PO₄</td>
<td>1PHZ</td>
<td>1</td>
<td>2.7/3.4 (Glu330)</td>
<td>(20)</td>
</tr>
<tr>
<td>rPheOH-Fe(III)</td>
<td>2PHM</td>
<td>3</td>
<td>2.7/3.4 (Glu330)</td>
<td>(20)</td>
</tr>
<tr>
<td>CvPheOH-Fe(III)</td>
<td>1LTV</td>
<td>2</td>
<td>2.2/2.5 (Glu186)</td>
<td>(14)</td>
</tr>
<tr>
<td>CvPheOH-Fe(III)·7,8-BH₂</td>
<td>1LTZ</td>
<td>2</td>
<td>2.2/2.2 (Glu186)</td>
<td>(14)</td>
</tr>
<tr>
<td>rTyrOH-Fe(III)</td>
<td>1TOH</td>
<td>3</td>
<td>2.1/2.6 (Glu376)</td>
<td>(25)</td>
</tr>
<tr>
<td>rTyr-Fe(III)·7,8-BH₂</td>
<td>2TOH</td>
<td>2</td>
<td>2.0/2.7 (Glu376)</td>
<td>(26)</td>
</tr>
<tr>
<td>hTrpOH-Fe(III)·7,8-BH₂</td>
<td>1MLW</td>
<td>3</td>
<td>2.4/3.4 (Glu317)</td>
<td>(7)</td>
</tr>
</tbody>
</table>

A recent X-ray structure of hPheOH determined with the natural cofactor and the substrate analogue THA (hPheOH-Fe(II)·BH₄·THA) revealed that the protein undergoes large conformational changes distributed throughout the entire molecule upon substrate binding. The structural changes involve the active site residues and trigger the Glu330 to change from a monodentate to a bidentate coordination of the active site iron. It has been suggested that this conformational change of Glu330 is necessary to provide space around the active site iron prior the formation of the putative catalytic oxyferryl species (23). Interestingly, the structures of rTyrOH-Fe(III) and the bacterial CvPheOH show a bidentate iron of the glutamic acid in the ferric forms of the enzyme (14, 21).
The bidentate coordination of the iron binding Glu184 of CvPheOH has been suggested to contribute to the tenfold higher activity of CvPheOH compared to the hPheOH (14).

### 1.7 Enzymatic mechanism of human phenylalanine hydroxylase

The crystal structures (discussed in chapter 1.5) of PheOH together with biochemical data have during the recent years increased our understanding of the enzymatic mechanism. The mechanism can be divided into two parts, first the generation of an oxidizing species, an activated oxygen (oxyferryl) and then the attack of this oxygen on the aromatic ring of L-Phe (23). The generation of the oxyferryl species is the rate-limiting step in tyrosine hydroxylase (31-33). A model of L-phenylalanine hydroxylation has been proposed by Andersen (23) (Figure 6).

During the enzymatic mechanism of PheOH the \textit{para} position hydrogen of the aromatic ring is transferred to the \textit{meta} position through hydrogen atom migration (NIH-shift). After the irreversible rate-limiting formation of the oxygenating intermediate the following steps are rapid (32). Consequently, measurements of kinetic parameters with amino acids of varying reactivity yield little information regarding the mechanism of hydroxylation and the molecular details of the NIH-shift are unknown. A hypothetical model has been suggested involving a Fe-O-L-Phe intermediate (Figure 7) (34).
Prior to catalysis, the active site iron is reduced by the cofactor BH$_4$ (Step 1) (35). During the reduction, the affinity for W1 and W2 is reduced and Glu330 is displaced (22). After the pre-reduction, reversible binding of BH$_4$ can occur (step 2). This will change the overall geometry of the active site. The side chain of Glu330 will change its conformation, coordinating the iron from a new position (22). Further conformational changes occur when L-Phe binds reversibly to the active site (step 3). The side chain of Glu330 adopts bidentate iron coordination and the position of the BH$_4$ is altered, allowing dioxygen binding at the position occupied by wat2 (23). After substrate binding, a putative Fe(II)-O-O-BH$_4$ intermediate is formed (step 4) (36, 37). Through heterolytic cleavage of the oxygen, an activated putative oxyferryl species is formed (step 5). This produces a molecule of 4a-OH-BH$_4$ and an activated oxygen intermediate. The details of the mechanism behind the dioxygen activation are still controversial (23). The activated oxygen (oxyferryl) can attack and hydroxylate the aromatic ring of L-Phe.
The hydroxylation proceeds mainly through the so-called NIH-shift (Figure 7). After hydroxylation the products are released (step 6).

(Figure 7) A suggested mechanism of the NIH-shift, the para position hydrogen of the phenylalanine ring is transferred to the meta position during the reaction (34). The details of this mechanism remain to be investigated.
2. The gram-negative, opportunistic pathogen

*Pseudomonas aeruginosa*

*Pseudomonas aeruginosa* is a gram-negative opportunistic bacterium that is noted for its ability to thrive in many ecological niches, from water and soil to animal and plant tissue. *Pseudomonas aeruginosa* is also noted for its resistance to many antibiotics and it causes clinical problems due to its ability to infect patients suffering from cystic fibrosis, cancer and burn wounds. *Pseudomonas aeruginosa* is one of the model organisms that is widely studied by scientists who are interested not only in its ability to cause disease and resist antibiotics, but also its metabolic capacity and environmental versatility. The genome sequence of *Pseudomonas aeruginosa* became recently publicly available (38).

2.1 The aromatic amino acid hydroxylating system of

*Pseudomonas aeruginosa*

*Pseudomonas aeruginosa* possesses in its *phh* operon genes related to the human PheOH and PCD/DCoH (Figure 8) (*phhA*, and *phhB* encoding the proteins PaPheOH/*phhA* and PaPCD/*phhB*, respectively) (15). The PaPheOH protein was formerly known as *phhA* and has been purified as a monomer with a deduced molecular weight of 30,288 Da (262 residues) (15). Sequence analysis of PaPheOH reveals that the protein has approximately 34% sequence identity to the catalytic domain of hPheOH, hTyrOH and hTrpOH (with 212 residues overlap) while the identity to the bacterial CvPheOH is 50% over a 131 residues interval (Figure 1).
The bacterial proteins lack the regulatory and tetramerization domains found in the mammalian proteins.

Figure 8. Organization of the structural genes within the *phh* operon of *Pseudomonas aeruginosa*. *PhhR* encodes a divergently transcribed regulatory protein, a member of the bacterial σ-factors, *phhA* encodes a phenylalanine hydroxylase (PaPheOH/PhhA), *phhB* encodes PaPCD/phhB, a carbinolamine dehydratase involved in pterin recycling and *phhC* encodes an aromatic aminotransferase.

The PaPCD/phhB protein (also discussed in chapter 2.2) has a deduced molecular weight of 13 333 Da (118 residues) (15) and shows a sequence identity of about 33% to the human analogue PCD/DCoH (103 residues interval). The bacterial protein forms homodimers in solution while the human analogue exists as homotetramers.

The physiological role of the *phh* operon of *Pseudomonas aeruginosa* is unknown. Expression of *phhA* and *phhB* is induced by the presence of L-phenylalanine when bacteria are grown on minimal medium (16). However, the primary route for tyrosine biosynthesis in Gram-negative bacteria is the widely distributed cyclohexadienyl dehydrogenase (39). Interestingly, PaPheOH is essential for the use of L-phenylalanine and L-tyrosine as a sole source of carbon in *Pseudomonas aeruginosa*. This suggests that the *phh* operon is responsible for the biosynthesis of some specialized compound starting from L-phenylalanine. The hypothesis is further supported by the fact that the
phh operon is induced although better carbon sources than L-phenylalanine (such as glucose) is supplied in the media (39).

The gene of the aromatic aminotransferase PhhC is located downstream of phhB in the phh operon (40). Assuming an operation of the phh operon in a catabolic mode would utilize the following steps (40): L-phenylalanine → L-tyrosine → 4-hydroxyphenylpyruvate.

Some of the genes that encode proteins involved in biopterin production and recycling in higher organisms have been identified in Pseudomonas aeruginosa. The biopterin cofactor is produced similar to the human system by the enzymes GTP cyclohydrolase I (NCBI accession no. AAG07441) and 6-pyruvoyl tetrahydrobiopterin synthase (NCBI accession no. AAG06054). So far no sepiapterin reductase or dihydropterin reductase has been reported.

2.2 PaPCD/phhB, a bacterial orthologue of PCD/DCoH

Pterin 4a-carbinolamine dehydratase (PaPCD/phhB) catalyzes the dehydration step in the regeneration of the tetrahydrobiopterin cofactor (Figure 4) (15, 41). In mammals, the homologue PCD/DCoH is a bifunctional protein also playing a regulatory role in the nucleus where it acts as a dimerization cofactor for the transcriptional activator HNF-1α. The mammalian protein forms homotetramers while PaPCD/ phhB exists only in a dimeric form [Song, 1999 #137, unpublished results].

Analysis of PaPCD/phhB using tyrosine autotrophy in Escherichia coli as a functional test revealed that the in vivo function of PaPheOH required the presence of PaPCD (16). Expression of PaPheOH without PaPCD induced toxic effects in Escherichia coli, probably due to nonenzymatic formation of 7-biopterin or related derivates (16, 42). The
PaPCD protein has a significant basal level of expression that is lacking for PaPheOH but both proteins are induced co-ordinately in the presence of either L-Phe or L-Tyr (16).

Translational lacZ reporter fusion experiments have indicated that PaPCD activates the phhA gene at the posttranscriptional level (16). Interestingly, this effect was also observed for the mammalian homologue, PCD/DCoH. Possible mechanisms for this activation have been suggested by Song et al. (16). The PaPCD protein may bind to PaPheOH mRNA and either protect the mRNA from degradation or enhance translational initiation. Another possibility is that PaPCD provides some catalytic- or stability enhancement of PaPheOH. Supporting the later theory is the result from immunoprecipitation experiments that demonstrated coprecipitation of PaPCD and PaPheOH (16).

When the structure of rPheOH with an intact regulatory domain was published, a structural similarity between the regulatory domain of hPheOH and PCD/DCoH was revealed (20). Since the PaPheOH protein is located in the same multigene operon as PaPCD in Pseudomonas aeruginosa this operon may have preceded the assembly of the two genes into a modular gene by exon shuffling (20).
3. Aims of the present study

The aim of the work presented in this thesis is the structural and biochemical characterization of a bacterial phenylalanine hydroxylase. The *Pseudomonas aeruginosa* phenylalanine hydroxylase was chosen as a model system on the basis of several arguments of which the most important were:

- The DNA sequence of PaPheOH encodes a phenylalanine hydroxylase that is one of the shortest known today, corresponding to the catalytic core of this group of enzymes.
- The *phh* operon that encodes the PaPheOH and PaPCD enzymes has an organization (discussed in chapter 2.1-2.2), which may have preceded the mammalian multi domain enzymes through exon shuffling (20). Moreover, a protein-protein complex composed of PaPheOH and PaPCD has been reported (16).

With the crystal structure of PaPheOH we can analyse to which degree the differences in amino acid sequence affect the structure of PaPheOH compared to hPheOH, rTyrOH, hTrpOH and CvPheOH. The X-ray structure of PaPheOH provides an opportunity to try to understand the catalytic activity and specificity from a structural point of view. We can also investigate the proposed interactions between PaPheOH and PaPCD since the structure of PaPCD was recently determined (U. Sauer, unpublished results). By using XAS to complement the X-ray studies we can obtain an independent view of active site environment and determine the coordination and oxidation state of the active site iron in the presence of substrate and/or cofactor.
4. X-ray crystallography

Today, two methods are used to obtain atomic resolution structures of macromolecules, X-ray crystallography and nuclear magnetic resonance (NMR). In this thesis, we have used X-ray crystallography complemented by XAS to investigate the atomic structure of PaPheOH. In the following chapter, I will briefly outline the technique.

4.1 Why we use X-rays and crystals

In organic molecules, the covalent bonding distances between atoms are usually around 1-2 Å. In order to study molecules at atomic resolution it is necessary to use radiation with a wavelength comparable to this distance. Electromagnetic radiation in this wavelength range is known as X-rays. X-rays interact with the electron cloud of atoms; however, the scattering information from an individual molecule is far too weak to be measured. Therefore, an absolute requirement for structure determination by X-ray crystallography is that the molecule of interest can be crystallized.

During crystal growth, units of the molecule (e.g. protein, DNA, RNA or an entire virus) are systematically incorporated into unit cells that form a three dimensional periodic lattice, the crystal. Each crystal contains a very large number of unit cells (about $10^{15}$), which means that the scattered X-rays from one unit cell can be amplified up to $10^{15}$ times as a result of the regular crystal packing. This makes it possible to record the signal onto an appropriate detector and determine the atomic structure of the molecule or particle in question.
4.2 Crystallization of proteins
The crystallization process is the most crucial and least understood part of the different steps that leads to X-ray structure determination of a molecule. Crystallization generally requires well-defined conditions that are molecule specific and impossible to predict prior to crystallization. It is very important that the sample is pure and homogenous. The growth of a crystal is induced in solution through a nucleation event. In order to obtain nucleation it is necessary that the system (protein solution) is in a supersaturated state. This is a labile condition where more of the protein is dissolved than theoretically possible, which means that the system is displaced from equilibrium so that the restoration requires precipitation of the protein, ideally under the formation of a crystal. A common way to identify supersaturated conditions is by vapour diffusion method. The sample is screened with a number of solutions of different compositions, often by using commercially available crystallization screens. In a typical experiment, the sample is mixed with an equal volume of the screening solution and placed on a cover slip. The cover slip is inverted and sealed over a container containing the screening solution ("hanging drop" technique) (Figure 9a). The drop is allowed to equilibrate against the screening solution by diffusion through the vapour phase. During equilibration, the volume of the drop will change affecting the concentration of the sample in the drop. In a typical experiment the well solution contains a higher concentration of the precipitant than the hanging drop. Equilibration will hence reduce the volume of the drop, resulting in a higher concentration of the protein solution. In a successful experiment, the sample becomes supersaturated and nucleation initiates crystal growth (Figure 9b). The time required to grow X-ray quality
crystals is not predictable and may take from a couple of days to several months.

After initial crystallization conditions are found they are refined and important parameters affecting crystal growth and diffraction quality are identified and optimized. Successful identification of crystallization conditions is considered one of the major bottlenecks within protein X-ray crystallography. Today, it is common that several hundreds different conditions are screened for each protein in order to identify crystallization conditions that yields crystals of good quality.

Figure 9. a) In a typical hanging drop vapour diffusion experiment; a drop of the sample is mixed with the screening solution on a cover slip before the sample is sealed over the screening solution in a closed container. b) Crystals of PaPheOH with a size of about 0.6 x 0.3 x 0.3 mm.

4.3 Data collection and analysis

Data collection refers to the process in which a crystal is subjected to X-rays and the resulting diffraction pattern is recorded on a two dimensional detector (Figure 10). When the X-ray beam travels through the crystal the electron cloud surrounding the atoms within the sample
will interact with the electromagnetic radiation (X-rays). Bragg diffraction occurs whenever the condition for constructive interference is satisfied (Bragg's law). During data collection, the crystal is exposed to the X-ray beam and a number of images are recorded on the detector. The crystal is rotated or oscillated with a fixed angle (typically 1 °) during each exposure. The number of images that are necessary to obtain a complete data set depends on the symmetry (space group) of the diffracted crystal. In a typical experiment, 90-180 ° of data is collected.

The aim of the data collection is to collect a complete data set, which means to record all possible Bragg reflections (spots). Often a unique reflection is recorded several times during the data collection, improving the accuracy of the measurement. After data collection, the diffraction spots are indexed and the space group and unit cell dimensions of the crystal are determined. During subsequent data processing, the spot profiles and the intensities are determined. Finally the intensities (I) are scaled and transformed into structure factor amplitudes (I \(\propto |F^2|\)). After scaling, several important features of the data are evaluated, usually in resolution shells. These include the completeness, the intensity distribution and the merging R-factor. Generally, data are collected to a completeness of more than 95%. Another important parameter is the R-factor, which describes the quality of the diffraction pattern. There are different types of R-factors and they all reflect the accuracy of the data. The R-factor should be as small as possible and for a data set of high quality it may be as low as 2%.
4.4 The initial model

After collection and initial analysis of the data one has to overcome another very crucial step, which is often referred to as "solving the phase problem". The phase angle associated with the structure factor amplitude must be determined for each of the recorded Bragg reflections. When the structure factor amplitude and phase angle for each Bragg reflection is known it will be possible to calculate (by Fourier transformation) the electron density distribution that gave rise to the initial scattering. The determined electron density will be used in later stages to build an atomic model that is able to reproduce the observed electron density. Today, there are several different ways to overcome the phase problem. This section will give a short overview of some important methods.

The "traditional" way of obtaining phase information is the so-called multiple isomorphous replacement (MIR) method. This method is
based on the comparison of several data sets, one collected from a native crystal and at least two data sets collected from crystals that have different strong scatterer bound to the protein. The strong scatterer is often an atom of a heavy metal compound that binds to a specific site of the protein and becomes part of the ordered array in the crystal. By subtracting the intensities of a crystal without any heavy atom (native data) from the corresponding intensities collected from crystals with heavy atom (derivate data) one is left with the scattering contribution from just the heavy atoms. This reduces the problem of determining the phase angle for many thousands atoms in the asymmetric unit to solving the phase angle only for a few heavy atoms. This can be accomplished by applying "Patterson methods". Once the heavy atom sub structure is solved for at least two heavy atom derivates one can determine the phase angle for the rest of the reflections.

Today, another technique for phase determination has increased in popularity: Multiple Anomalous Dispersion (MAD), in which data at several different wavelengths are collected around the absorption edge, e.g. peak-, inflection- and remote data. Similar to MIR, this method depends on small differences of the diffraction intensities that electron dense atoms cause. These atoms absorb the X-rays at an element specific wavelength, the absorption edge (also discussed in chapter 4.1). By collecting diffraction data close to the absorption edge, the anomalous contribution of these atoms affects the intensities of the Bragg reflections. By analyzing the anomalous contribution to the intensities, it is possible to determine the phase angles of the Bragg reflections. An important variant of the MAD technique is single anomalous dispersion
(SAD) in which a single data set is collected, typically at the peak wavelength.

Another widely used method to obtain phase information is molecular replacement (MR). In molecular replacement, the 3D structure of a protein or subdomain with high sequence identity to the protein of interest is used as a model for a MR search. The search is usually done in two-step; typically, the model is first subjected to a rotation search. During the rotational search the model is rotated and the correlation between the model and the experimental data is evaluated. After the rotational search promising rotational positions are tested in the translational search. Now the correct position of the molecule in the asymmetric unit is identified. With a correct solution of the rotational and translational search the phase angle for each Bragg reflection can be calculated.

4.5 Model building and structure refinement

After the phase angle has been determined for each of the structure factor amplitudes (Bragg reflections), the electron density that gave rise to the initial scattering can be calculated by Fourier transformation. The electron density is visualized using computer graphics (Figure 11a) which allows a model to be built into the density (Figure 11b). In general, this is done manually with the help of computer graphics using specialized programs and the known amino acid sequence. For high-resolution structures it may be possible to use computer programs for automatic model building.
Figure 11. a) Experimental electron density at 1.6 Å resolution (√σ-weighted 2F<sub>c</sub>-F<sub>c</sub> map at 1.0 σ). b) Model built into the electron density. In the centre of the picture is an iron atom visible. A glutamic acid, a histidine and a water molecule coordinate the iron.

The initial model is often a crude and incomplete representation of the real structure. To improve the initial model, several iterative cycles of model building and refinement are necessary. The purpose of the refinement is to minimize the errors and to optimize the model in order to obtain the best possible correlation between the calculated structure factors (F<sub>calc</sub>) based on the model and the experimentally observed structure factors (F<sub>obs</sub>). In a model building cycle, the main- and side chain of each amino acid are fitted into the electron density. After inspection and rebuilding, the model is energy minimized and a new electron density is determined based on phases calculated from the improved model. The model is rebuilt into the new density and the building cycle is repeated.

To judge the quality of the model, a R-factor is calculated. Two different variants of the R-factor are in use, R and R<sub>free</sub>. Both describe the correlation between the structure factors calculated from the model and the experimental data, R<sub>free</sub> is considered more important since it is
unbiased. To calculate $R_{\text{free}}$, a set of reflections is set aside (test set) from the reflections that are used for refinement (working set). The progress of the refinement is monitored by the calculation of an $R_{\text{free}}$ value based only on the excluded reflections. The equation for the $R_{\text{free}}$ factor is given below.

$$R(\%) = 100 \times \frac{\sum |F_{\text{test}}| - |F_{\text{calc}}|}{\sum |F_{\text{test}}|}$$

In this equation, $|F_{\text{test}}|$ are the experimentally observed structure factor amplitudes (from the test set) and $|F_{\text{calc}}|$ are the calculated structure factor amplitudes based on our model. For well-determined X-ray structures the R factor is around 20% and the $R_{\text{free}}$ a few percent higher. The R-values can be as low as 5-10% for an exceptional model determined to high resolution.
5. X-ray Absorption Spectroscopy

X-ray Absorption Spectroscopy (XAS) is an experimental technique that can provide complementary information to an X-ray structure about bond distances, coordination number as well as the oxidation state and geometry of a specific element. XAS is a technique that is useful for samples in any physical state and it uses a much lower X-ray dose during data collection compared to X-ray crystallography. The low dose excludes radiation damage.

5.1 X-ray absorption spectroscopy

When a sample is irradiated with electromagnetic radiation in the X-ray region, the electrons within the sample will interact with the electric field of the radiation (also discussed in chapter 4.2). Either the radiation will be elastically scattered by these electrons (a phenomena used in X-ray crystallography), or the radiation will be absorbed and lead to an excitation of the core electrons of the sample. The absorption of energy and the excitation of the core electrons is a process that can be studied by X-ray absorption spectroscopy (XAS). XAS measurements are performed around the element specific absorption edge. The absorption edge is the energy, expressed in keV, when the absorption of the radiation dramatically increases (Figure 12). Each such edge occurs when the energy or wavelength of the incident photons is sufficient to lift a core electron into a higher orbital or expel it from the absorbing atom. The energy when the photoelectron is released is defined as the threshold
energy of the absorption edge ($E_0$). The threshold energy normally corresponds to the first inflection point on the low energy side of the absorption edge.

![Graph showing absorption edge](image)

Figure 12. A typical absorption edge recorded for the active site iron of PaPheOH.

5.2 Interpretation of XAS spectra

The absorption spectra are often divided into several regions, which have their unique characteristics and can provide information about different aspects of the chemical environment surrounding the absorbing atom. The X-ray absorption near edge structure (XANES) region is located within about ±10 eV of the main absorption edge. The pre-edge features can provide information about coordination geometry while the edge position is influenced by the oxidation state of the absorbing atom.

The near-edge X-ray absorption fine structure (NEXAFS) region extends from about 10 to 50 eV above the edge. This region contains information about the atoms within the first coordination shell around the
absorber as well as information about tightly bound atoms beyond the first coordination shell.

The extended X-ray absorption fine structure (EXAFS) region extends from energies of about 50 eV up to more than 1000 eV above the absorption edge. The weak oscillations observed in this region provide information about the binding distances between the absorber and atoms on a distance of up to 6 Å. For a successful interpretation of the EXAFS spectra it is necessary to have a good theoretical model whose back-calculated EXAFS spectra can be compared and correlated to the experimentally observed data (Figure 13). Often, several different coordination models must be tested and evaluated before model that agrees with the experimental data is found.

Figure 13. The experimental EXAFS spectra of PaPheOH and its correlation to the theoretical EXAFS spectra of an evaluated model.
With this study, we initiate our characterization of PaPheOH and reported the cloning, purification and crystallization of PaPheOH. The protein crystallized in space group P61 with unit cell dimensions of a=b=210.509 Å c=100.699 Å for a native data set that diffracted to 2.0 Å resolution. Matthews constant indicated 6-12 molecules in the asymmetric unit. The relatively high number of molecules in the asymmetric unit in combination with the low sequence identity to the used model proteins (hPheOH and rTyrOH) complicated the molecular replacement trials and no clear solutions could be identified. In an attempt to obtain phase information we collected a three wavelength MAD data set around the active site Fe k-edge. The data was evaluated with SOLVE but no clear solution could be identified (43). In an attempt to circumvent the weak anomalous signal displayed by the Fe k-edge data we collected a three wavelengths MAD data set around the selenium k-edge of crystals containing protein where the Met was replaced with Se-Met. The anomalous signal of the Se-Met data was sufficient to obtain phase information in a SAD phasing trial using CNS (44). The structure of PaPheOH is presented in paper III.
In this report we describe X-ray absorption spectroscopy (XAS) measurements of the active site iron of PaPheOH. Samples of PaPheOH and PaPheOH with cofactor (BH$_4$) and/or substrate (L-Phe) were subjected to XANES/EXAFS spectroscopy at the DESY synchrotron beamline X1, Germany. The XANES spectroscopy indicates that the active site iron is coordinated by six ligands forming a distorted octahedral orientation. Analysis of the oxidation state indicates that the active site iron is in a catalytically active ferrous form (FeII). Oxidation with dopamine or H$_2$O$_2$ yields the ferric (Fe(III)) form of the protein. In samples of PaPheOH with L-Phe and/or BH$_4$ the intensity of the pre-edge peak reflects a higher degree of symmetry. The increase of symmetry might be due to the formation of hydrogen bonds between the BH$_4$ cofactor and the water molecules that coordinates the iron.

EXAFS spectroscopy agrees with a model in which the iron is coordinated by two histidines (His121, His126), one glutamic acid (Glu166) and additional water molecules. Based on our data we suggest that Glu166 coordinates the iron in a bidentate geometry. The coordination of the active site iron is unaffected by addition of cofactor and/or substrate. In a sample of PaPheOH with the BH$_4$ cofactor the Deby-Waller parameters increased for the second and third coordination shells. This indicates structural modification for structure at these shells and/or an increase in the disorder. This is consistent with X-ray studies of hPheOH.
This manuscript describes the high resolution X-ray structure of PaPheOH. The overall structure of the catalytic core of PaPheOH is composed of a basket-like arrangement of helices, strands and loops. The active site cleft is located in the centre of the catalytic domain basket. The iron is positioned at the bottom of the active site cleft where it is coordinated by the side chains of His121, His126, Glu166 and two additional water molecules. Consistent with our model that derived from the XAS study (paper II), Glu166 coordinates the iron in a bidentate fashion. The overall coordination of the iron resembles the coordination observed in hPheOH-Fe(II)-BH$_4$·THA structure as well as the structures of CvPheOH-Fe(III) and CvPheOH-Fe(III)-7,8-BH$_2$ (14, 23).

Many of the residues assigned to pterin binding in CvPheOH and hPheOH are found at a different position in PaPheOH. The so-called "pterin binding loop" (residue 81-86) is located at a distance of 5-6 Å from the expected cofactor binding position. We suggest that the unfavourable position of the pterin-binding loop is the main cause for the low activity observed for PaPheOH, although the bidentate coordination of Glu166 would suggest high activity.

The expected substrate-binding site of PaPheOH is similar to the corresponding region of CvPheOH and hPheOH. High atomic B-factors and weak electron density for the loop region (residue 106-118) surrounding the expected substrate binding site of PaPheOH suggest flexibility in this region. This observation is consistent with results reported from CvPheOH (14).
Skeletal muscle $\alpha$-actinin is one of the key organizers within the Z-disc of striated muscle. In an attempt to structurally characterize this multidomain protein we focused our attention to the N-terminal located actin binding domain (ABD). The ABD of $\alpha$-actinin comprises about 260 residues and is composed of two calponin homology domains (CH-domains). The crystallization of ABD was complicated and the growth of high quality crystals was sporadic and unpredictable. Interestingly, the crystallization attempts of native ABD were unsuccessful while a E37K mutant produced crystals. The protein crystallized in space group P2$_1$ with the unit cell dimensions of $a=101.90$ b=$38.40$ c=$154.90$ Å $\beta=109.20^\circ$. Matthews's constant indicates four molecules in the asymmetric unit. Molecular replacement trials using the structures of related proteins failed to produce a solution. Due to the highly irreproducible crystallization behaviour of the protein we decided to use the Multiple Anomalous Dispersion (MAD) technique on selenomethionine (Se-Met) labelled protein crystals to obtain phase information. Crystals of Se-Met labelled ABD grew in clusters to a size of about 15 $\mu$m x 15 $\mu$m x 100 $\mu$m. We used the microfocusing beamline ID-13 at the ESRF synchrotron in Grenoble, France to collect a Single Anomalous Dispersion (SAD) data set. Phasing trials using SOLVE and CNS were unsuccessful (43, 44). Analysis of the data revealed that the anomalous signal displayed a fast decrease during data collection. This decrease we attribute to radiation damage of the crystal
10. The 2.0 Å X-ray structure of the actin binding domain of α-actinin

Fredrik Ekström, Stefan Bäckström, Gunter Stier, Kristina Djinović Carugo and Uwe H. Sauer

Recently, we were able to solve the 2.0 Å X-ray structure of the actin binding domain (ABD) of human α-actinin by molecular replacement methods using coordinates provided by K. Djinović Carugo (unpublished results). The asymmetric unit contains four molecules of ABD, corresponding to a solvent content of 47%. The following description is based on the current not fully refined model. The ABD comprising residue 20-243 folds into two domains composed of 5 α-helices each (Figure 14). They are linked through a loop connecting helix 5 and helix 6.

Figure 14. Ribbon diagram of ABD with the secondary structure elements labelled sequentially (H1-H10). The helices fold into two distinct CH domains (CH1 in and CH2) connected through a loop between helix 5 and helix 6. Conserved residues that have been implicated in direct interaction with actin are coloured red (45, 46).
The electron density covers Trp165, a residue located in the CH2 domain.

The electron density is weak for the N-terminal CH1 domain whereas the C-terminal CH2 domain has a well-defined electron density (Figure 15).

The two CH domains adopt a compact, closed conformation with extensive intramolecular contacts between the two domains. Actin binding domains has been suggested to undergo a conformational switch between a closed and an open state. X-ray structures of utrophin and dystrophin represent the open state (47, 48), while the compact arrangement of the CH domains of α-actinin is similar to the closed states found in fimbrin and plectin (49, 50).

The refinement of actin binding domain of α-actinin is still in progress. We will soon present the fully refined structure and the structural analysis of the ABD domain of α-actinin.
11. Conclusions

The findings described in this thesis can be summarized as follows:

The X-ray structure of PaPheOH shows that the side-chains of His121, His126, Glu166 and two water molecules octahedrally coordinate the active site iron. The glutamic acid coordinates the iron with bidentate geometry that corresponds to coordination found in CvPheOH and the substrate activated form of hPheOH.

Our XAS model of the active site environment agrees and support the coordination observed in the X-ray structure of PaPheOH.

The pterin binding loop of PaPheOH-Fe(III) is located at a distance of 5-6 Å from the expected pterin binding site. The unfavourable position of the pterin binding loop may provide a structural explanation for the low specific activity of PaPheOH compared to CvPheOH and PaPheOH.

The two CH domains of β-actinin are composed of 5 α-helices each. The CH domains are similar to the other members of the CH domain super family. Together they fold into a closed conformation with extensive intramolecular contacts between the two domains.
12. Future perspectives

The work presented in this thesis provides a structural explanation for the low activity observed for PaPheOH. However, there are many details of cofactor and substrate binding that remains to be investigated for the bacterial enzyme. Of particular interest is the structure of a complex between PaPheOH and the cofactor or a similar derivative. This has been the focus of much of my work, but I have not been able to determine such a structure. In the future, we will continue to subject PaPheOH to crystallization trials in an anaerobic environment. We hope it will provide structural information about PaPheOH in complex with cofactor and/or substrate analogues. Such an approach has proven successful for the human enzyme.

The proposed interaction between PaPheOH and PaPCD remains to be investigated. Size exclusion chromatography experiments have not been able to reproduce the suggested PaPheOH-PaPCD complex (results not published). Possibly, some component necessary for the complex formation is absent in our system. This remains to be investigated with techniques such as Biacore and/or NMR.
13. Acknowledgements

Först av allt vill jag tacka min handledare Uwe för all erfarenhet du gett mig och för alla diskussioner om stort och litet. My collaborator, Gunter I really want to thank you for your enthusiasm, for your "know how" and for providing me with interesting proteins to crystallize. I'm looking forward to take a hike in the mountains soon.

Stefan "black box" Bäckström, jag är så otroligt tacksam för allt stöd i alla lägen. Tack för att du har stått ut med alla dumma frågor och för att du visat mig hur en struktur ska lönas! Nu när vi båda disputerat får vi kanske tid för vår grottur! Andreas, dina råd om livet och program har alltid varit värdefulla. Ingen kan leta fel som du! Sen är du ju också ganska trevlig att umgås med... Anders, tack för att du alltid ställt upp och diskuterat prylar, dina kunskaper på området är oöverträffade! Erik, alltid lika positiv och rolig att prata med. Synd att jag slutar just som jag börjar förstå din skånska.... (Ja, ja, jag vet att du kommer från Blekinge). Liz, tack för din entusiasm som gör att allt känns mycket lättare, uppskattas! Shenghua,thanks for all advice and assistance with MR. And good luck with your plantings, both your crystals and your garden.... Christin, min favorit slavdrivare! Ingen kan piska som du! Tack för att du alltid frågar hur det är och sen dessutom har ork att lyssna... Tobias, for a realistic and interesting view on life in general and science in particular... Good luck... Ulrika... Du är rolig du... Det kommer att gå bra för dig! Marek vill jag tacka för att du alltid har hjälpt mig med trilskandes datorer. Jeanette, dina script är super... Snart är det din tur... Karina och Anders, så synd att jag åker just när kristallografin i Umeå får en sån förstärkning. Lycka till!

Jag vill också tacka Linda Fahlen, Linda Öberg, Christina, Talal, Josefin och alla andra vid UCMP. Teresa, ditt stöd under skrivandet av avhandlingen har varit otroligt värdefullt, snart är det din tur, det kommer att gå bra!

Min tutor och blivande medarbetare Susanne Lundberg, tack för att du alltid har funnits där och stöttat. Det ska bli så otroligt roligt att jobba med dig i framtiden.
Sen har vi alla mina fantastiska vänner, ni är otroliga...


Jag vill också tacka mina älskade föräldrar för att ni gett mig alla möjligheter man kan önska. Tack för att ni står ut och stöder alla mer eller mindre dumma projekt jag hittar på. Utan er uppmuntran och stöd hade jag aldrig nått fram.

“*If you go on hammering away at a problem, it seems to get tired, lies down and lets you catch it*”

*W.L. Bragg*